

IN THE CLAIMS

Please amend the following claims:

sub 01
1 1. A method for document classification comprising:
2 analyzing textual and graphical properties of a previously unclassified electronic
3 document using text data and image data of the entire electronic document;
4 generating a classification of the document based on the textual and graphical
5 properties; and
6 storing the electronic document in a pre-existing directory structure based on the
7 classification.

01
1 2. The method defined in Claim 1 where the directory structure comprises a
2 hierarchy of documents mirroring in a similar fashion an organization in a pre-existing
3 memory storing documents.

1 3. The method defined in Claim 2 when the pre-existing memory comprises
2 a hard disk.

sub 02
1 4. The method of claim 1, wherein analyzing textual properties of an
2 electronic document comprises:
3 determining characteristic words of the document;
4 determining a frequency for each characteristic word; and

5 building a frequency table based on the frequency associated with each
6 characteristic word.

1 5. The method of claim 1, wherein analyzing graphical properties of an
2 electronic document comprises:

3 determining a point set corresponding to the electronic document, wherein points
4 of the point set correspond to end points of lines;

5 determining a density of points within the point set;

6 generating a document profile based, at least in part, on the density of points
7 within the point set.

1 6. The method of claim 1, wherein generating a classification of the
2 document based on the textual and graphical properties comprises combining results from
3 the textual and graphical analysis using a Borda Count.

1 7. The method defined in Claim 1 further comprising building the pre-
2 existing directory structure by building hierarchy of documents based on a user's hard
3 drive.

5w
03 } 1 8. The method defined in Claim 1 further comprising building the pre-
2 existing directory structure by extracting graphical and text features from documents in a
3 directory-based memory to obtain a

1 9. A software product including a machine-readable medium having stored
2 thereon sequences of instructions, which, when executed by a processor, cause the
3 processor to:

4 analyze textual and graphical properties of a previously unclassified electronic
5 document using text data and image data of the entire electronic document;

6 generate a classification of the document based on the textual and graphical
7 properties; and

8 store the electronic document in a pre-existing directory structure based on the
9 classification.

1 10. The machine-readable medium of claim 9, wherein the sequences of
2 instructions that cause the processor to analyze textual properties of an electronic
3 document further comprise sequences of instructions that cause the processor to:

4 determine characteristic words of the document;

5 determine a frequency for each characteristic word; and

6 build a frequency table based on the frequency associated with each characteristic
7 word.

1 11. The machine-readable medium of claim 9, wherein the sequences of
2 instructions that cause the processor to analyze graphical properties of an electronic
3 document further comprise sequences of instructions that cause the processor to:

4 determine a point set corresponding to the electronic document, wherein points of
5 the point set correspond to end points of lines;

6 determine a density of points within the point set;
7 generate a document profile based, at least in part, on the density of points within
8 the point set.

1 12. The machine-readable medium of claim 9, wherein the sequences of
2 instructions that cause the processor to generate a classification of the document based on
3 the textual and graphical properties further comprises sequences of instructions that cause
4 the processor to combine results from the textual and graphical analysis using a Borda
5 Count.

1 13. A method for document classification comprising:
2 analyzing documents in a pre-existing document directory structure to determine a
3 document classification profile of the pre-existing document directory structure;
4 generating a mirror directory structure based on the pre-existing document
5 directory structure;
6 receiving a previously unclassified electronic document;
7 analyzing textual and graphical properties of the electronic document using text
8 data and image data of the entire electronic document; and
9 placing the electronic document in the mirror directory structure based on the
10 document classification profile of the pre-existing document directory structure, results of
11 textual analysis of the document, and results of graphical analysis of the document.

1 14. The method of claim 13, wherein analyzing the pre-existing document
2 directory structure further comprises:

3 recursively descending the pre-existing document directory structure;
4 generating a list of directories in the pre-existing document directory structure;
5 examining files in directories of the pre-existing document directory structure to
6 determine content; and
7 examining the content of the files to determine document classification profile of
8 the directories in the pre-existing document directory structure.

1 15. The method of claim 13 wherein the pre-existing document directory
2 structure is a hard disk directory structure.

1 16. The method of claim 13 wherein generating a mirror directory structure
2 based on the pre-existing document directory structure comprises generating a document
3 directory structure having a set of directories and relationships equivalent to the pre-
4 existing document directory structure.

1
1 17. The method of claim 13 wherein placing the electronic document in the
2 mirror directory structure comprises:
3 determining a primary directory in the pre-existing document directory structure
4 in which the document should be placed based on the document classification profile of
5 the pre-existing document directory structure; and
6 storing the document in a primary corresponding directory in the mirror directory
7 structure that corresponds to the primary directory in the pre-existing document directory
8 structure.

1 18. The method of claim 17 further comprising:
2 determining a secondary directory in the pre-existing document directory in which
3 the document should be placed based on the document classification profile of the pre-
4 existing document directory structure; and
5 storing the document in a corresponding secondary directory in the mirror
6 directory structure that corresponds to the secondary directory in the pre-existing
7 document directory structure.

al
Sub
B3
1 19. A computer-readable medium having stored thereon sequences of
2 instructions which, when executed by a processor, cause the processor to:
3 analyze a pre-existing document directory structure to determine a document
4 classification profile of the pre-existing document directory structure;
5 generate a mirror directory structure based on the pre-existing document directory
6 structure;
7 receive a previously unclassified electronic document;
8 analyze textual and graphical properties of the electronic document using text data
9 and image data of the entire electronic document; and
10 place the electronic document in the mirror directory structure based on the
11 document classification profile of the pre-existing document directory structure.


1 20. The computer-readable medium of claim 19, wherein the sequences of
2 instructions that cause the processor to analyze a pre-existing document directory
3 structure to determine an organization of the pre-existing document directory structure
4 further comprise sequences of instructions that cause the processor to:

5 recursively descending the pre-existing document directory structure;
6 generating a list of directories in the pre-existing document directory structure;
7 examining files in directories of the pre-existing document directory structure to
8 determine content; and
9 examining the content of the files to determine the organization of the directories
10 in the pre-existing document directory structure.

a 1 21. The computer-readable medium of claim 19, wherein the sequences of
2 instructions that cause the processor to generate a mirror directory structure further
3 comprise sequences of instructions that cause the processor to generate a document
4 directory structure having a set of directories and relationships equivalent to the pre-
5 existing document directory structure.

1 22. The computer-readable medium of claim 19, wherein the sequences of
2 instructions that cause the processor to place a document in the mirror directory structure
3 further comprise sequences of instructions that cause the processor to:
4 determine a primary directory in the pre-existing document directory structure in
5 which the document should be placed based on the document classification profile of the
6 pre-existing document directory structure; and
7 store the document in a primary corresponding directory in the mirror directory
8 structure that corresponds to the primary directory in the pre-existing document directory
9 structure.

1 23. The computer-readable medium of claim 22 further comprising sequences
2 of instructions that cause the processor to:
3 determine a secondary directory in the pre-existing document directory in which
4 the document should be placed based on the document classification profile of the pre-
5 existing document directory structure; and
6 store the document in a corresponding secondary directory in the mirror directory
7 structure that corresponds to the secondary directory in the pre-existing document
8 directory structure.

 1 24. An apparatus comprising:
2 means for analyzing a pre-existing document directory structure to determine
3 document classification profile of the pre-existing document directory structure;
4 means for generating a mirror directory structure based on the pre-existing
5 document directory structure;
6 means for receiving a previously unclassified electronic;
7 means for analyzing textual and graphical properties of the electronic document
8 using text data and image data of the entire electronic document; and
9 means for placing the electronic document in the mirror directory structure based
10 on the document classification profile of the pre-existing document directory structure.

1 25. The apparatus of claim 24, wherein means for analyzing the pre-existing
2 document directory structure further comprises:
3 means for recursively descending the pre-existing document directory structure;

4 means for generating a list of directories in the pre-existing document directory
5 structure;
6 means for examining files in directories of the pre-existing document directory
7 structure to determine content; and
8 means for examining the content of the files to determine document classification
9 profile of the directories in the pre-existing document directory structure.

a 1 26. The apparatus of claim 24, wherein means for generating a mirror
2 directory structure comprises means for generating a document directory structure having
3 a set of directories and relationships equivalent to the pre-existing document directory
4 structure.

sub 37 1 ~~27. The apparatus of claim 24, wherein means for placing a document in the~~
2 ~~mirror directory structure comprises:~~
3 ~~means for determining a primary directory in the pre-existing document directory~~
4 ~~structure in which the document should be placed based on the document classification~~
5 ~~profile of the pre-existing document directory structure; and~~
6 ~~means for storing the document in a primary corresponding directory in the mirror~~
7 ~~directory structure that corresponds to the primary directory in the pre-existing document~~
8 ~~directory structure.~~

1 28. The apparatus of claim 27 further comprising:

2 means for determining a secondary directory in the pre-existing document
3 directory in which the document should be placed based on the document classification
4 profile of the pre-existing document directory structure; and
5 means for storing the document in a corresponding secondary directory in the
6 mirror directory structure that corresponds to the secondary directory in the pre-existing
7 document directory structure.

1 29. A document processing system comprising:
2 a document scanning device;
3 a document storage device coupled to the document scanning device, wherein the
4 document storage device is organized as a document directory structure having multiple
5 directories, and further wherein the document storage device has a mirror directory
6 structure having multiple directories organized based on the document directory
7 structure; and
8 a processor coupled to the document scanning device and to the document storage
9 device, wherein the processor analyzes content of a document scanned by the document
10 scanning device to determine a directory in the document directory structure in which the
11 document should be placed and stores the document in a corresponding directory in the
12 mirror directory structure.

1 30. The document processing system of claim 29 wherein the processor is
2 operable to determine a secondary directory in the document directory structure in which
3 the document should be placed and to store the document in a corresponding secondary
4 directory in the mirror directory structure.

31. The document processing system of claim 29 wherein the processor

- 2 analyzes files stored in the document directory structure to determine content and
3 generates a document classification profile of the document directory structure based on
4 the analysis.

32. The document processing system of claim 29 wherein the document is

- 2 analyzed based on image and textual content.
-